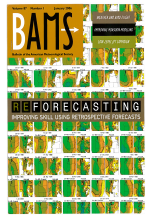
Using reforecasts to improve probabilistic weather predictions

Tom Hamill, Jeff Whitaker, and Gary Bates
ESRL Physical Sciences Division



What is a reforecast?

A "hindcast," a large set of weather or climate predictions *using the same model and data assimilation system that is used to make the real-time forecasts*.

Why generate reforecasts?

Weather and climate models often have large systematic errors. To provide useful decision support to customers, forecasts must be as accurate and reliable as possible. Reforecasts facilitate the "calibration" of weather forecasts, the post-processing to statistically adjust today's forecast using the discrepancies between past forecasts and observations.

What makes this different from MOS?

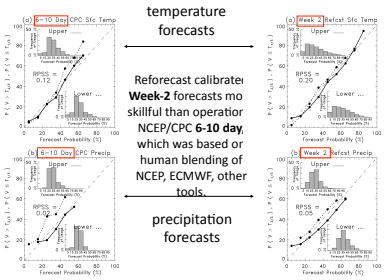
"MOS" is Model Output Statistics, a similar technique used by NOAA's Meteorological Development Lab for producing statistically corrected forecasts based on NCEP's models. Whereas MOS uses whatever training data is available from NCEP, sometimes a very short training data set given the ever-changing NCEP models, our purpose was to explore what can be done when a *very large training data set* from a frozen model is available. Also, we have developed more sophisticated statistical techniques for post-processing than were used in MOS. The large training sample size from reforecasts is shown to be especially helpful in statistically adjusting forecasts for rare events (e.g. heavy precipitation) and forecasts at long leads.

What reforecast data sets have been computed?

First, a reforecast from a T62L28 (~250 km grid spacing) version of the NCEP GFS from 1998. 15-member ensemble forecasts have been computed out to 15 days lead for every day from 1979 to present. Since generating this data set, NCEP/EMC's Climate Forecast System Reanalysis and Reforecast (CFSRR) has been produced, configured using lessons learned from our 1st-generation reforecast ECMWF produces reforecasts in real time following our work. And we expect to do a next-generation reforecast for NCEP's current operational Global Ensemble System.

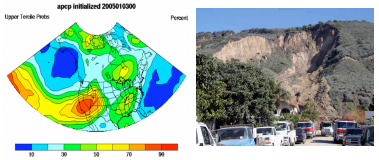
Weather-climate applications

Improving forecasts at long leads is a difficult problem of determining what forecast signal is there hidden in the noise created by the chaotic growth of initial condition errors and model error. We successfully demonstrated that reforecast-based products were more skillful than the operational 6-10 day and week-2 forecast produced by the NCEP Climate Prediction Center (CPC). As a result, CPC now uses our reforecast products.



Example: La Conchita, CA mudslide, 1/12/05

notice very high probabilities of above-normal precipitation forecast over California



Selected references

Hamill, T. M., J. S. Whitaker, and X. Wei, 2004: Ensemble re-forecasting: improving medium-range forecast skill using retrospective forecasts. *Mon. Wea. Rev.*, **132**, 1434-1447.

Hamill, T. M., J. S. Whitaker, and S. L. Mullen, 2006:Reforecasts, an important dataset for improving weather predictions. *Bull. Amer. Meteor. Soc.*, **87**,33-46.

Hamill, T. M., and J. S. Whitaker, 2006: Probabilistic quantitative precipitation forecasts based on reforecast analogs: theory and application. *Mon. Wea. Rev.*,**134**, 3209-3229.

Wilks, D. S., and T. M. Hamill, 2007: Comparison of ensemble-MOS methods using GFS reforecasts. *Mon. Wea. Rev.*, **135**, 2379-2390.

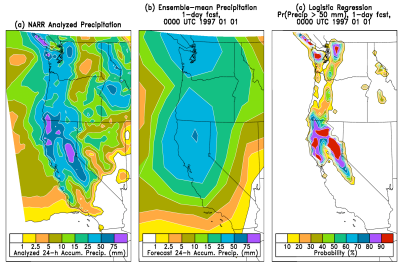
Hagedorn, R, T. M. Hamill, and J. S. Whitaker, 2008: Probabilistic forecast calibration using ECMWF and GFS ensemble forecasts. Part I: 2-meter temperature. *Mon. Wea. Rev.*, **136**, 2608-2619.

Hamill, T. M., R. Hagedorn, and J. S. Whitaker, 2008: Probabilistic forecast calibration using ECMWF and GFS ensemble forecasts. Part II: precipitation. *Mon. Wea. Rev.*, **136**, 2620-2632.

Hagedorn, R., R. Buizza, T. M. Hamill, M. Leutbecher, and T. N. Palmer, 2010: Comparing TIGGE multi-model forecasts with reforecast-calibrated ECMWF ensemble forecasts. Mon. Wea. Rev., in review.

Heavy-precipitation applications

Producing accurate, reliable forecasts of heavy precipitation can be difficult for many reasons. Raw forecast model output is commonly biased, and the forecasts may not be able to be computed at high-enough resolution to predict the small-scale features. With a large training sample size available from reforecasts, it is possible to make statistical adjustments to heavy precipitation forecasts, statistically downscaling the forecasts and dramatically improving forecast skill. Because of the success with reforecasts, the NCEP Hydrometeorological Prediction Center uses reforecast-based products to improve their quantitative precipitation forecasts.



Example of the statistical downscaling and calibration that is possible using reforecasts. Left panel shows the analyzed precipitation; middle panel shows the ensemble-mean precipitation forecast, and right panel the calibrated probability of greater than 50 mm precipitation using the ensemble-mean data as a predictor.

What comes next?

We will generate a new reforecast data set with a current-generation version NCEP global ensemble system. We will: (a) make it available for post-processing research at ESRL, EMC, MDL, and universities. Here at ESRL, we will to use the data set to develop prototype applications for long-lead severe weather threats, aviation-related products, and renewable energy applications (solar output, wind-energy potential) (b) use it to define how NCEP should compute reforecasts operationally in future years. With a reforecast available every day, for example, we can determine how much worse the statistical post-processing performs for crucial applications with a sample every second day, or every third day. (c) Collaborate with hydrological colleagues, e.g. NOAA's Office of Hydrologic Development, to develop calibrated probabilistic ensemble streamflow guidance using the reforecasts. (d) Test whether the data set is useful for removing bias in data assimilation.

Will reforecasts provide as much benefit with newer models?

To explore this, we worked with colleagues at ECMWF to test the benefit of reforecasting with their ensemble prediction system. We found that (1) there was still a very large benefit to applying the reforecast technology to their system; and (2) our old, 1998 model with calibration provided through reforecasts provided more skillful probabilistic guidance than their ensemble prediction system without calibration. (3) The impact of the large training sample size was greater for rare events (heavy precipitation) than it was for more common events (light precipitation).

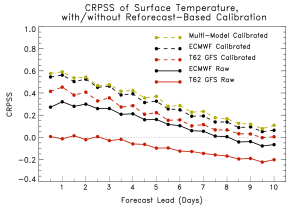
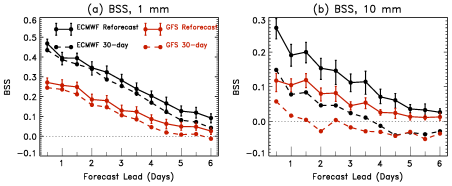


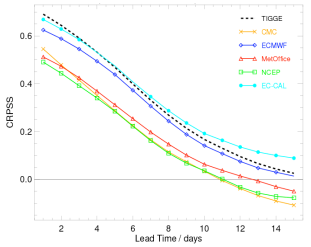
illustration of the skill of surface temperature forecasts from ECMWF and GFS reforecast systems, before and after reforecast calibration.



Precipitation Brier skill scores for (a) light precipitation (1 mm/day) and (b) heavier precipitation (10 mm/day), before and after reforecast calibration. Notice the greater impact of the reforecast calibration for the 10-mm event.

Multi-model ensembles instead?

ECMWF, with reforecast calibration, produced as good or better temperature forecasts than were obtained through a multi-model combination and calibration using a shorter, 30-day training data set.



Calibrated ECMWF reforecast temperature skill (light blue) was generally as good or better than individual models' skill and the 30-day calibrated multi-model combination.